Monthly Mean Diurnal Cycles in Surface Temperatures over Land for Global Climate Studies

ALEXANDER IGNATOV

UCAR Visiting Scientist, NOAA/NESDIS, Office of Research and Applications, Climate Research and Applications Division, Washington, D.C.

GARIK GUTMAN

NOAA/NESDIS, Office of Research and Applications, Climate Research and Applications Division, Washington, D.C.

(Manuscript received 21 October 1997, in final form 30 June 1998)

ABSTRACT

Monthly mean diurnal cycles (MDCs) of surface temperatures over land, represented in 3-h universal time intervals, have been analyzed. Satellite near-global data from the International Satellite Cloud Climatology Project (ISCCP) with a $(280\ \text{km})^2$ resolution (C-2 product) are available for seven individual years and as a climatology derived thereof. Surface 19-yr climatologies on ground and air temperatures, separately for all-sky and clear-sky conditions, matched with the ISCCP data, are employed to better understand satellite-derived MDCs.

The MDCs have been converted to local solar time, refined to a regular 1-h time grid using cubic splines, and subjected to principal component analysis. The first two modes approximate MDCs in air and ground–satellite temperatures with rmse's of about $\sigma = 0.5°$ and 1°C, respectively, and these accuracies are improved by 20%–35% if the third mode is added. This suggests that two to three temperature measurements during the day allow reconstruction of the full MDC. In the case of two modes, optimal observation times are close to the occurrence of minimum and maximum temperatures, $T_{\min}$ and $T_{\max}$. The authors provide an empirical algorithm for reconstructing the full MDC using $T_{\min}$ and $T_{\max}$, and estimate its accuracy. In the analyzed match-up dataset, the statistical structure of ground temperature for all-sky conditions most closely resembles that of the ISCCP derived temperature. The results are potentially useful for climate- and global-scale studies and applications.

1. Introduction

The information on the diurnal cycle of land surface temperature, T_S, is important for many meteorological, climatological, and remote sensing applications. These include but are not limited to validation and/or checking the consistency of the numerical weather and climate models (e.g., Cao et al. 1992); estimating surface energy and water balances, thermal inertia, and soil moisture (e.g., Idso et al. 1975; Price 1977, 1980; Carlson et al. 1981; Wetzel et al. 1984; Tarpley 1994; Seguin et al. 1994); monitoring crops (e.g., Casellas and Sobrino 1989); reconstructing the full diurnal cycle from only a few temperature measurements during the day and, in particular, from the two extrema ($T_{\min}$ and $T_{\max}$); removal of biases induced by changing observation times in standard meteorological (Schaal and Dale 1977; Karl et al. 1986) and satellite data (Privette et al. 1995; Gutman and Ignatov 1995); and advancement of global climate change studies (Karl et al. 1993). The T_S diurnal cycle depends on many factors, which are aggregated into two major components: the surface energy balance, dependent on insolation, surface properties, and atmospheric variables; and the surface thermal inertia, dependent on the type of soil, its moisture content, and vegetation cover (e.g., Price 1977, 1980; Oke 1987). The complex character of the processes at the surface–atmosphere interface, and the fact that many input parameters are not easily obtained, make it difficult to formulate a practical physical model of the diurnal cycle. An alternative way to describe it is through a statistical analysis of empirical data, without explicit description of the involved physical mechanisms.

The diurnal march at any point on an individual day (synoptic scale) may be highly volatile. It becomes more deterministic and predictable as data are averaged over larger time- and space scales. This paper uses monthly mean diurnal cycles (MDCs) collected in a wide variety of geographical areas, which makes its results oriented mostly at the climate- and global-scale applications, for instance, in climate modeling. Global satellite data regularly gridded in space and time appear to be most suit-

Corresponding author address: Dr. Garik Gutman, NOAA/NESDIS E/RA1, 4700 Silver Road, Stop 4700, Washington, DC 20233-9910.
E-mail: ggutman@nesdis.noaa.gov

able for this purpose. Here, we analyze the International Satellite Cloud Climatology Project (ISCCP) C-2 data (Rossow and Schiffer 1991), which are monthly (280 km)^2 averages for seven individual years, and their 7-yr mean. For quality control and better understanding of the satellite data, we use standard meteorological data collected at 75 stations in the former Soviet Union and kindly provided to us by Dr. P. Groisman (National Climatic Data Center) as 19-yr climatological MDCs of air and ground temperatures for all-sky conditions and separately for its clear-sky subset.

We employ principal component analysis and find that the first two modes approximate the MDCs with root-mean-square errors (rmse's) of about 0.5°C for air and 1°C for ground and satellite data, while adding the third mode increases the accuracy by ≈20%–35%. Alternatively, an equivalent number of temperature measurements during the day can be used instead. We analyze when those are to be taken in order to best approximate the MDCs and find that in the case of two modes the optimal times are close to the occurrence of daily minimum and maximum. Once these findings are verified in the various experimental data types, the mapped values of a few principal components (or temperatures) provide the data for future analyses. Their time series and/or maps can be used to monitor interannual change and/or spatial variability at the surface (e.g., Karl et al. 1993). Another potential application is checking the consistency between GCM diurnal cycles (e.g. Cao et al. 1992) and the empirical EOF statistics. Yet another subject for research would be associating temperature ranges with surface types. Often in the literature, this type of discussion is conducted for individual daily temperature cycles, whereas this paper provides the approach for climatic analysis.

Direct comparison of satellite and station temperatures is hindered by their different natures, scales, and measurement technology. Therefore, we intercompare their statistical structures rather than the absolute values and find that satellite data in the analyzed match-up dataset most closely resemble all-sky ground station temperature, whereas all other station temperatures form a statistical pattern quite different from satellite data. This empirical fact somewhat defies our intuition but being unambiguously observed in the data, allows a few different interpretations, and we discuss some of them. Yet it may be sample biased and needs further checking.

2. Data

a. Satellite data

Data on the T_S diurnal cycle can be routinely collected from a few geostationary satellites, each operationally observing a certain part of the globe. Aggregating data from all available platforms and mapping them into a regular space–time grid is a tedious task. It was accomplished in the framework of the ISCCP, which had produced a multiyear dataset of satellite-derived surface temperatures, T_{sat}, for the July 1983–December 1990 period. Top-of-the-atmosphere clear-sky radiances, obtained at the cloud detection stage, were corrected for the atmospheric effect and used for the derivation of T_{sat}, assuming surface emissivity to be 1. The variable ''mean surface temperature for clear-sky composite'' (a by-product of the cloud-oriented ISCCP), available from the ISCCP Monthly Cloud Product C-2 on CD-ROM (ISCCP 1992), is a statistical summary of the satellite observations. Data are presented as monthly MDCs in eight 3-h intervals, from 0000 through 2400 UTC, mapped globally in an equal-area (280 km)^2 projection. In addition to individual year data, we have derived a 7-yr monthly climatology of ISCCP T_{sat} (based on 1984–90) for comparison with the 19-yr climatology of the former Soviet Union (fUSSR) in situ measurements described in the next section.

FIG. 1. Coverage by the ISCCP data (circles) and the 75 fUSSR stations (dots).

Only land grid cells within a ±60° latitude–longitude belt (not including coasts) were analyzed, which amounts to 14 940 (12 × 1245) MDCs per year (Fig. 1; the data gap over India and north of it, including a large part of the fUSSR, is caused by the unavailability of data from the Indian satellite, *INSAT*, for the ISCCP). We hereafter refer to this sample as the ''global'' sample. No data modification such as quality control, screening, or smoothing was applied except procedures described in section 2c.

The ISCCP data are advantageous in that they provide a complete diurnal cycle on a near-global, uniform, and internally consistent basis. However, the climate community is unaccustomed to thinking in terms of T_{sat}. The major reasons are probably the short time record of ISCCP T_{sat} (now extending to one decade), and the expected clear-sky bias in T_{sat} (the ISCCP T_{sat} has been derived from infrared measurements in cloud-free pixels only, in contrast to all-sky measurements in traditional meteorology). Eventually, longer records of T_{sat} will be aggregated as future satellite data are collected. This paper compares T_{sat} with the conventional temperature measurements in an attempt to improve its value for the use in many climate and meteorological applications.

b. Station data

Four 19-yr monthly climatologies[1] of the in situ diurnal cycles of air temperature at a shelter level, and ground temperature[2] at the surface, at 223 fUSSR stations were used. This study considers two types of climatologies: all sky (T_{aa} and T_{ga}), and clear sky (T_{ac} and T_{gc}, derived from a subsample corresponding to cloud-free conditions). Our particular objective is to check on the concern expressed by some climatologists that satellite retrievals, being made under cloud-free conditions, may be biased toward clear-sky conditions (e.g., Groisman and Zhai 1995). The MDCs at each station are presented in eight 3-h intervals, from 0000 to 2400 UTC, consistent with ISCCP data. More details on the data can be found in Groisman et al. (1997).

Merging the satellite and station climatologies yields only 75 points (indicated in Fig. 1 by dots), out of 223, largely because of the data gap over Siberia. In what follows, we refer to this subsample as the ''stations'' sample, containing 900 realizations (75 $\times$ 12), and consider it along with the full ISCCP global sample.

c. Recalculating to a regular grid local time

Both satellite and station data are given in UTC, whereas the thermal processes at the surface, driven by the sun, depend upon local solar time. The first step therefore was to recalculate UTC to local time (LT) as follows: LT = UT $-$ λ/15, where λ is longitude. After conversion, an interpolation procedure was used to generate data on a regular grid in a local timescale and to fill in missing data. The time grid has been refined at this stage from 3 h UTC to 1 h LT (we stress that this does not change the number of parameters in a diurnal cycle to be investigated later using principal component analysis). Linear interpolation was tested but rejected, in particular because it systematically underestimates the T_S maxima, especially in cases when only a few observations during the day are available, with the maximum often missed (the minima are flatter and therefore less subject to this error). This problem is greatly alleviated when cubic splines are used (e.g., Meinguet 1979). The splines are run in two-dimensional space (hour–month) with a special treatment at the boundaries (December–January, 24 LT–1 LT). To accommodate the derivatives, four points (12 h) were added on both ends of the daily cycle and two months on both ends of the annual cycle.

3. Principal component analysis (PCA)

Since insolation, the primary driving force for the thermal regime at the surface, is a periodic process, the diurnal cycle in T_S has been often approximated by trigonometric functions and Fourier series (e.g., Price 1977; Minnis and Harrison 1984; Cautenet et al. 1986; Lagouarde and Brunet 1993; Cracknell and Xue 1996). Keeping the minimum possible number of terms in a series is crucial for rectifying signal from noise. To that end, principal component analysis (PCA) allows description of the variability within a dataset with a minimal number of linear degrees of freedom, through using the eigenvectors of the covariance matrix that form the natural internal basis of the system (Preisendorfer and Mobley 1988). Any other externally imposed basis, including the above-mentioned trigonometric representations, requires at least as many or more terms to be kept to describe the experimental data with a compatible accuracy.

Here, we follow Kondragunta and Gruber (1994) and Cairns (1995), who pioneered application of PCA to the analysis of diurnal cycles in ISCCP data. An arbitrary MDC, of any temperature parameter, converted to local time and interpolated to a 1-h interval, is a 24-dimension vector $\mathbf{T}_S \equiv \{T_{S,1}, \ldots, T_{S,24}\}$, which is represented as a linear combination of the average, $\overline{\mathbf{T}}_S$, and a few specific (per unit α) diurnal cycles (or EOF), $\mathbf{e}_{S,n}$:

$$\mathbf{T}_S = \overline{\mathbf{T}_S} + \sum_{n=1}^{N} \alpha_{S,n} \cdot \mathbf{e}_{S,n}. \quad (1)$$

Equation (1) is a compressed representation of five different equations corresponding to five temperature parameters. The symbol ''*S*'' denotes either of the subscripts ''aa,'' ''ac,'' ''ga,'' ''gc'' (first subscript for *a*ir or *g*round, second for *a*ll or *c*lear), or ''sat.'' The 24-dimension vectors of $\overline{\mathbf{T}}_S$ and $\mathbf{e}_{S,n}$, which do not change in space and time, are calculated over the ensembles of 14 940 for the global and of 900 for the station samples. The full MDC is then succinctly described, according to (1), by N numbers—the weights (or principal components, PC) $\alpha_{S,n}$, which, in contrast to $\overline{\mathbf{T}}_S$ and $\mathbf{e}_{S,n}$, are location (station) and month specific.

Figure 2 shows the average and the first three specific diurnal cycles. Two statistics of $\mathbf{T}_{sat}$ (station and global) allow checking the stability and reproducibility of the PCA results with respect to the sample. The general shapes of $\mathbf{e}_{sat,n}$ for the two samples compare favorably, whereas the average vectors, $\overline{\mathbf{T}}_{sat}$, differ as would be expected, the regional sample being biased with respect to the global. Additional analysis has shown that, despite this bias, the station $\overline{\mathbf{T}}_S$ and $\mathbf{e}_{S,n}$ can be utilized to describe the variability within the global dataset with al-

[1] In fact, the base period differs slightly for different stations. Most stations extend their records from 1966 through 1983 or 1984, and five stations have data through 1986.

[2] Ground temperature is not as easy and straightforward to measure as the air temperature. Also, its use in representing large areas is problematic, given the diversity of surface types within a climatic grid cell. This parameter is not among the routine station measurements over large parts of the world. We refer the reader to the paper by Groisman et al. (1997) for a comprehensive description and discussion of the measurement procedure at the Russian meteorological stations. It was that procedure that had been used to collect the data used in the present study.

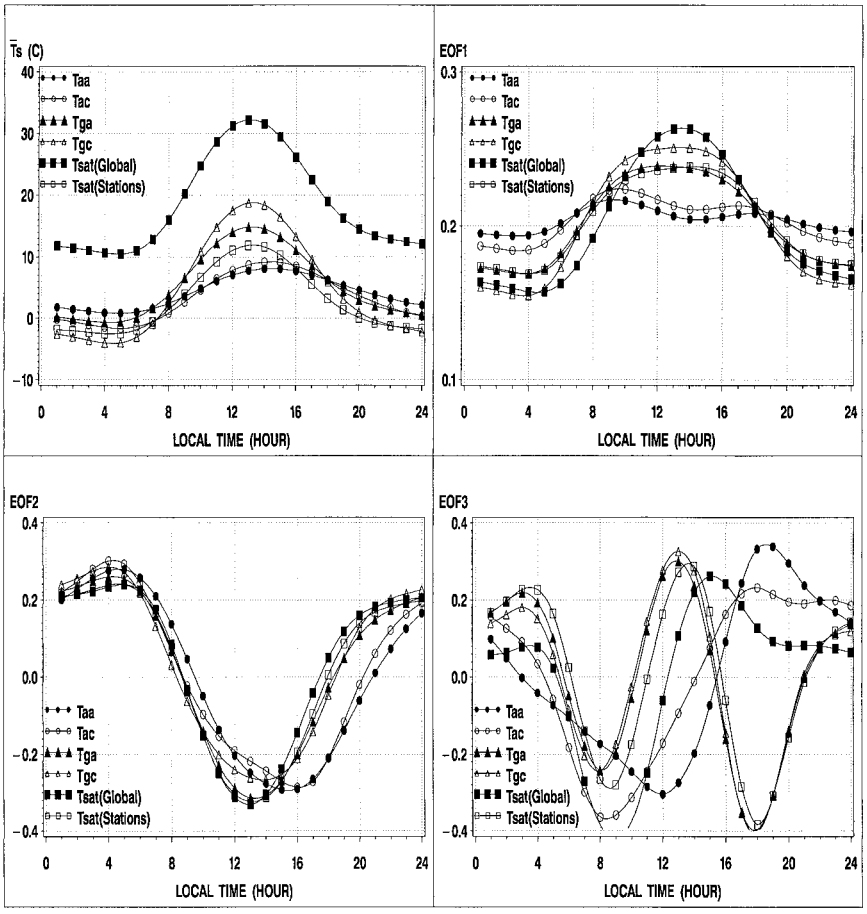

FIG. 2. Mean, $\overline{\mathbf{T}}_S$, and the first three empirical orthogonal functions: $\mathbf{e}_{S,1}$ (EOF1), $\mathbf{e}_{S,2}$ (EOF2), and $\mathbf{e}_{S,3}$ (EOF3) of T_{aa}, T_{ac}, T_{ga}, T_{gc}, and T_{sat} for the station, and of T_{sat} separately for the global sample.

TABLE 1. Residual rmse (°C) of approximating monthly mean diurnal cycles in different temperature parameters with different number of PCA modes (percent explained variance is shown in brackets).

Temperature parameter	Model			
	$\mathbf{T} = \overline{\mathbf{T}}$	$\mathbf{T} = \overline{\mathbf{T}} + \alpha_1\mathbf{e}_1$	$\mathbf{T} = \overline{\mathbf{T}} + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2$	$\mathbf{T} = \overline{\mathbf{T}} + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3$
T_{aa}	13.4	1.2	0.5	0.3
	(0)	(99.2)	(99.8)	(99.9)
T_{ac}	15.2	1.7	0.9	0.7
	(0)	(98.8)	(99.7)	(99.8)
T_{ga}	15.7	2.2	0.7	0.5
	(0)	(98.0)	(99.8)	(99.9)
T_{gc}	18.6	2.6	1.2	0.8
	(0)	(98.1)	(99.6)	(99.8)
T_{sat} (station)	14.8	2.1	0.9	0.6
	(0)	(98.0)	(99.7)	(99.8)
T_{sat} (global)	12.7	3.8	1.1	0.7
	(0)	(91.3)	(99.3)	(99.7)

most no loss in accuracy, and the global statistics works comparably well on the station sample.

Table 1 shows rmse's of approximating MDCs with different number of modes, along with the explained percent variance, and Fig. 3 gives a perspective of their diurnal distribution (one can arrive at the numbers in Table 1 by averaging respective data in Fig. 3 over the day, in root-mean-squared sense). No pronounced spatial or seasonal biases in the approximation error by either number of modes has been noticed. The first mode describes MDCs with rmse's of ≈1.2°C and ≈2.2°C for air and ground/satellite temperatures, respectively, in the station sample, and ≈4°C for satellite temperatures in the global sample. Adding the second mode reduces errors ≈2–3.5 times down to ≈0.5° and ≈1.0°C, respectively, explaining >99% variance for all temperature parameters. Further analysis in the paper mostly concentrates on two-parameter representation of the MDCs and does not explore potential accuracy improvement by ≈20%–35% if the third mode is added.

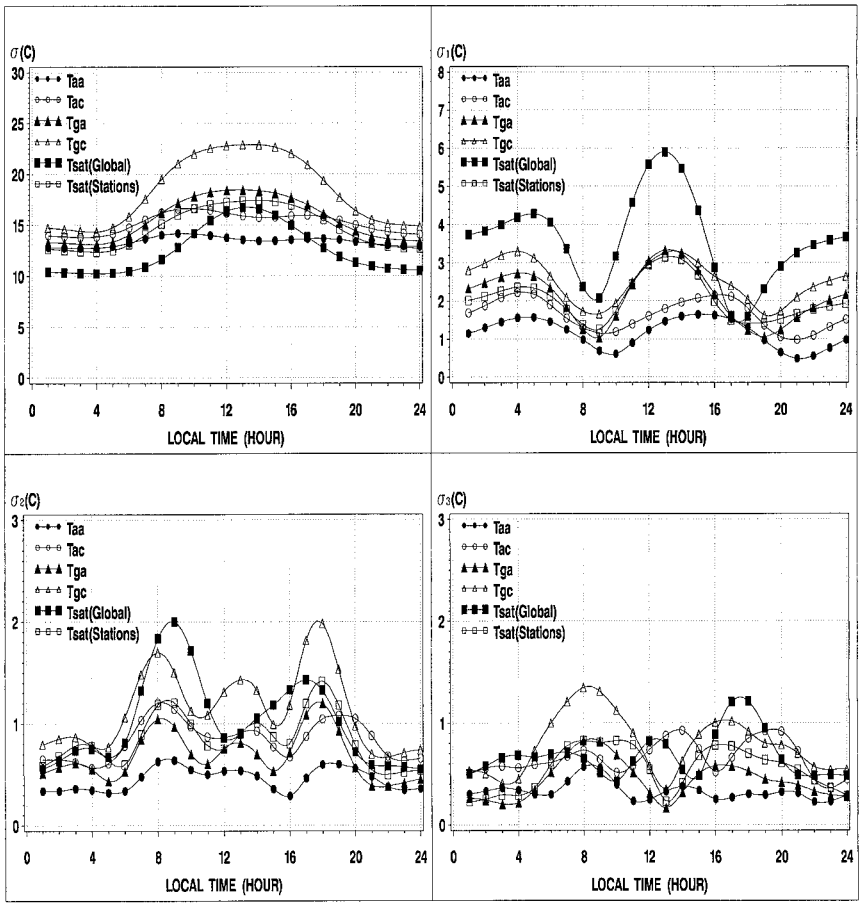

FIG. 3. Rmse of approximating diurnal cycles with different number of modes: $\mathbf{T}_S = \overline{\mathbf{T}}_S\,(\sigma)$, $\mathbf{T}_S = \overline{\mathbf{T}}_S + \alpha_{S,1} \cdot \mathbf{e}_{S,1}\,(\sigma_1)$, $\mathbf{T}_S = \overline{\mathbf{T}}_S + \alpha_{S,1} \cdot \mathbf{e}_{S,1} + \alpha_{S,2} \cdot \mathbf{e}_{S,2}\,(\sigma_2)$, and $\mathbf{T}_S = \overline{\mathbf{T}}_S + \alpha_{S,1} \cdot \mathbf{e}_{S,1} + \alpha_{S,2} \cdot \mathbf{e}_{S,2} + \alpha_{S,3} \cdot \mathbf{e}_{S,3}\,(\sigma_3)$.

4. Discussion

a. Comparison of different temperatures

To better understand the nature of different temperatures, we compare their statistical structures (shown in Fig. 2) rather than absolute temperature values that depend upon definitions and observational scales, and may be subject to systematic errors. Some sources of these errors (for instance, due to a calibration offset or to a violation of the blackbody assumption) are expected to have no pronounced diurnal variability but may depend upon the surface type [vegetation cover, topography, presence of water bodies within the $(280\ \text{km})^2$ grid cell]. Thus, absolute values of T_{sat} are subject to bigger uncertainties and therefore are expected to be less reliable than the shape of MDC and relative/differential factors, such as, for instance, diurnal range, in which some of these errors presumably cancel out.

Two ground temperatures, T_{ga} and T_{gc}, are consistent by the general shape of the mean and specific diurnal cycles, and in particular, by the times of occurrence of the extrema. The two air temperatures,[3] on the other hand, form a group that stands apart from these two. The diurnal march of air temperature lags the ground temperature by 1–3 h because of the energy flow (from the surface to the air in daytime, and from the air to the surface at night) and has smaller range.

The statistics of T_{sat} better resemble ground temperatures. In midlatitude Russia, of which the 75 stations are representative, the vegetation is senescent, absent, or covered by snow during a large part of the year (recall that the analysis was performed for all months together). This may introduce a sample-specific bias toward bare

[3] The local minimum in EOF1 of air temperatures around 1400 LT is consistent with the local minimum in the σ curve in Fig. 3, because EOF1 explains up to 98% of the variability (see Table 1). This means that the station-to-station and interseasonal variability (not air temperature itself, which has no minimum) decreases in the midday. At this point of analysis, we have no explanation of this effect. More in depth analysis is needed to determine if this phenomenon is a real feature, or if it is an artifact of data processing.

soil in the ''stations'' sample, possibly affecting the closeness of satellite and ground temperatures. Over dense vegetation, the satellite hardly sees the surface through vegetation so that T_{sat} is expected to better compare with air temperature. This problem needs further in-depth analysis.

Another important implication of Fig. 2 is that in the group of ground/satellite temperatures, T_{sat} stands closer to T_{ga}, not T_{gc}, as suggested by Groisman and Zhai (1995). A possible explanation is that although satellite retrievals are made over the pixels that are classified as cloud free, those pixels are still influenced by the ambient and preexisting average cloudiness.[4] The ISCCP clear-sky composites used here may thus have cloud effects built in, making T_{sat} more representative of all-sky rather than clear-sky conditions. The remaining discrepancies between $\overline{\mathbf{T}}_{sat}$ and $\overline{\mathbf{T}}_{ga}$ are attributed to the differences in spatial scales and sampling, to errors in satellite data (for instance, due to residual cloud contamination), and to the assumption of blackbody (unity emissivity).

b. Optimal observation times

If the MDC is characterized by two parameters, then a pair of temperature measurements taken at times h_1 and h_2, $T_S(h_1)$ and $T_S(h_2)$, can be used to predict temperature $T_S(h)$ at any arbitrary time h:[5]

$$T_S(h) = \beta_{S,0}(h; h_1, h_2) + \beta_{S,1}(h; h_1, h_2)T_S(h_1) + \beta_{S,2}(h; h_1, h_2)T_S(h_2) + \epsilon_S(h; h_1, h_2). \quad (2)$$

The accuracy of (2) depends on the observational times, h_1 and h_2. We define optimal times as those bringing an effective error (norm of the error) of approximating the diurnal cycle to a minimum: $(h_{1,opt}, h_{2,opt}) = \text{argmin}\|\boldsymbol{\epsilon}_S\|$. The norm may be defined, for instance, as a sum of squares (least squares problem, L2 metrics) or as a maximum possible error (minimax problem, L∞ metrics):

$$\|\boldsymbol{\epsilon}_S(h_1, h_2)\|_{L2} = \sqrt{\sum_{h=1}^{24} \epsilon_S^2(h; h_1, h_2)/24};$$
$$\|\boldsymbol{\epsilon}_S(h_1, h_2)\|_{L\infty} = \max_h |\epsilon_S(h; h_1, h_2)|. \quad (3)$$

Figure 4 shows $\|\boldsymbol{\epsilon}_S\|_{L2}$ for all possible combinations of h_i, h_j $(i, j = 1, \ldots, 24)$ for three temperature parameters S = ''aa,'' ''ga,'' and ''sat,'' the latter separately for the station and global samples. (The patterns for respective clear-sky temperatures closely resemble those for all skies and are not shown. Our analysis has shown that major conclusions hold if $\|\boldsymbol{\epsilon}_S\|_{L\infty}$ is used instead of $\|\boldsymbol{\epsilon}_S\|_{L2}$.) Best accuracy is achieved when $h_{1,opt} \approx 1200 \ldots 1400$ LT for ground and satellite temperatures (for air temperatures, $h_{1,opt} \approx 1400 \ldots 1600$ LT), and $h_{2,opt} \approx 0200 \ldots 0500$ LT. Although errors increase as the observational times deviate from optimal, the flat and wide minima of $\|\boldsymbol{\epsilon}\|_{L2}$ make the result of retrieval insensitive to the small shifts around these optima.

c. Description of MDCs using T_{min} and T_{max} instead of PCs

Interestingly, the optimal times are close to the occurrence of temperature diurnal extrema, T_{min} and T_{max}, whose observations are routinely taken at meteorological stations and from NOAA polar orbiting satellites. This suggests representing the MDC as[6,7]

$$\mathbf{T}_S = T_{S,min} \cdot \mathbf{f}_{S,1} + T_{S,max} \cdot \mathbf{f}_{S,2} + \boldsymbol{\epsilon}_S. \quad (4)$$

Figure 5 shows the 24-dimension vectors, $\mathbf{f}_{S,1}$ and $\mathbf{f}_{S,2}$, and the respective rmse vectors, $\sqrt{\langle\boldsymbol{\epsilon}_S^2\rangle}$. They were determined using multivariate linear regression analysis (where minimum and maximum, out of 24 temperatures, have been used in lieu of the actual extrema). Similarly to $\mathbf{e}_{S,n}$, the basic functions $\mathbf{f}_i$ cluster into two groups, one including ground/satellite and second air temperatures, with a remarkable consistency within each group. But in contrast to $\mathbf{e}_{S,n}$, $\mathbf{f}_{S,1}$ and $\mathbf{f}_{S,2}$ are not orthogonal anymore.

Equation (4) is advantageous in that it offers a direct method of calculating the full diurnal cycle through measurements of T_{min} and T_{max}. Its accuracy is worse compared to the two-mode fit by (1) (cf. σ_2 in Fig. 3 with the rmse in Fig. 5) because no representation with a given number of terms fits data better than PCA. Additionally, some error results from using integer-hour approximations of the extrema from spline-smoothed 3-h data, in place of their actual values, which are unknown. This error is expected to be random, thus influencing chiefly the accuracy of (4), not its coefficients, β. This preliminary conclusion, however, should be validated using refined measurements of extrema and full diurnal cycle.

To compare satellite and station data with their ap-

[4] Note that the clear-sky station data are influenced by the ambient and preexisting cloudiness, similar to the clear-sky satellite data. The definitions of ''clear sky'' for the satellite and station observation are different, as are the scales. Detailed analysis of this question is beyond the scope of this paper and needs further investigation.

[5] If $N > 2$ measurements during a day are available, one has a choice of solving a system of N equations, either with N $(\alpha_1, \ldots, \alpha_N)$ or with fewer $n < N$ $(\alpha_1, \ldots, \alpha_n$; an overdetermined system) unknowns, α standing for the respective principal components.

[6] Note that an average vector in Eq. (4), $\mathbf{f}_0$, is dropped. Its addition showed no significant reduction in the residual error, $\boldsymbol{\epsilon}_S$.

[7] Diurnal range ΔT and daily mean temperature T_{mean} can be used in Eq. (4) in place of T_{min} and T_{max}. Our estimates show that the results are close, since one pair of temperatures is easily calculated from another, by a simple linear transformation: $\Delta T = T_{max} - T_{min}$; $T_{mean} = T_{min} + \gamma\Delta T$. Interestingly, according to our regression analysis, γ differs from the customarily used value of 0.5: $\gamma \approx 0.46$ for air temperatures and $\gamma \approx 0.38$ for ground/satellite temperatures.

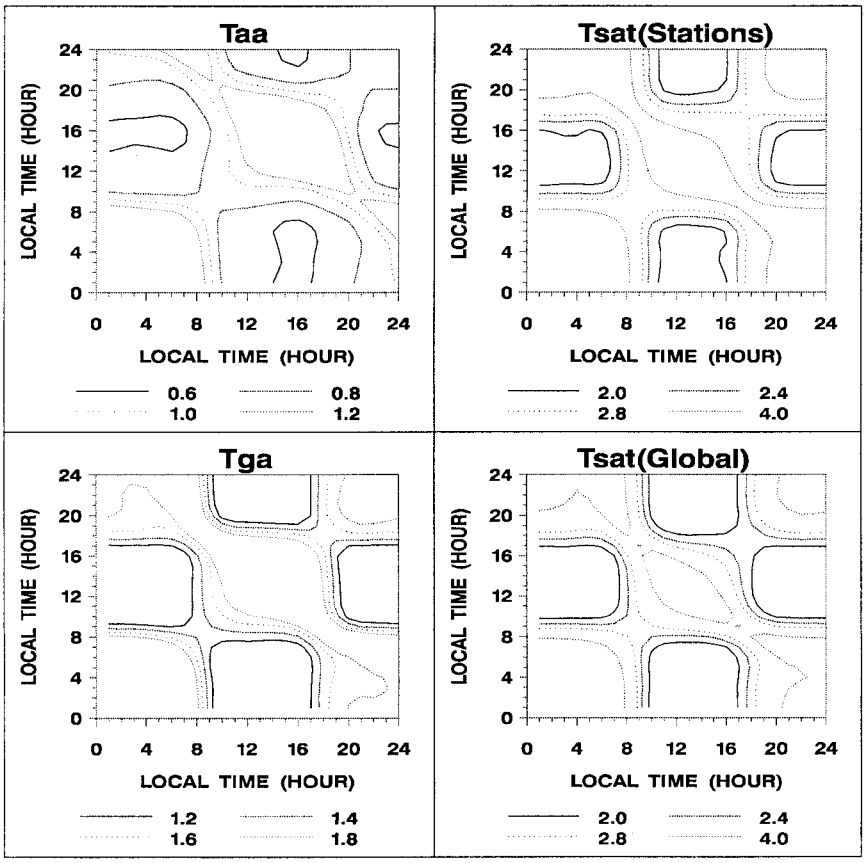

FIG. 4. Rmse of regressions $T_S = \beta_0 + \beta_1 T_S(h_1) + \beta_2 T_S(h_2)$. Minima indicate optimal times for observations.

proximations through $T_{\min}$ and $T_{\max}$, we selected the station Vladivostok, one of the few with data extending beyond 1983 and thus overlapping with the ISCCP data for the period of July 1983–August 1986. Daily data for this station were electronically accessed at the National Climatic Data Center and processed into monthly means of all-sky air and ground temperatures for individual years. Figure 6 shows 3-hourly satellite and station temperatures (dots) along with their approximations (lines) using $T_{\min}$ and $T_{\max}$. In cases of missing observations, the proposed method allows reconstructing the full MDC from only a few data points; for example, in June 1984 where only three observations, out of eight, are available. The approximation curves in Fig. 6, calculated using $\mathbf{f}_{S,1}$ and $\mathbf{f}_{S,2}$ derived from climatological statistics,[8] provide a good fit. The disrepancies between MDCs for T_{sat} and T_{ga} during the transition springtime are probably due to the fact that coastal station data are compared to a (250 km)^2 area including mountains.

Not only does the example in Fig. 6 illustrate the validity of the model given by (4), but it also shows a potential for the thermal mapping and interannual monitoring. Since the surface thermal regime is fully characterized by more than one parameter, be it PCs or a few temperature observations with the number depending upon the required accuracy,[9] a system for mapping/monitoring should use its complete set, not just one daily measurement. Seasonal and interannual variability, easily traced in Fig. 6, will be analyzed in future studies.

5. Conclusions

Monthly mean diurnal cycles (MDCs) of conventional and satellite temperatures over land have been

[8] Individual year MDCs are expected to be more volatile due to lesser averaging applied, and their statistics may deviate from the climatological one. At this point, the average and specific diurnal cycles were calculated for different years and found to be remarkably reproducible with no detectable deviation from climatological statistics.

[9] This result emphasizes the importance of including nighttime temperatures in research datasets from NOAA polar orbiters, directed at studying land surfaces. Present datasets (e.g., James and Kalluri 1994; Gutman et al. 1995) contain only daytime observations, making reconstruction of the diurnal temperature cycle infeasible.

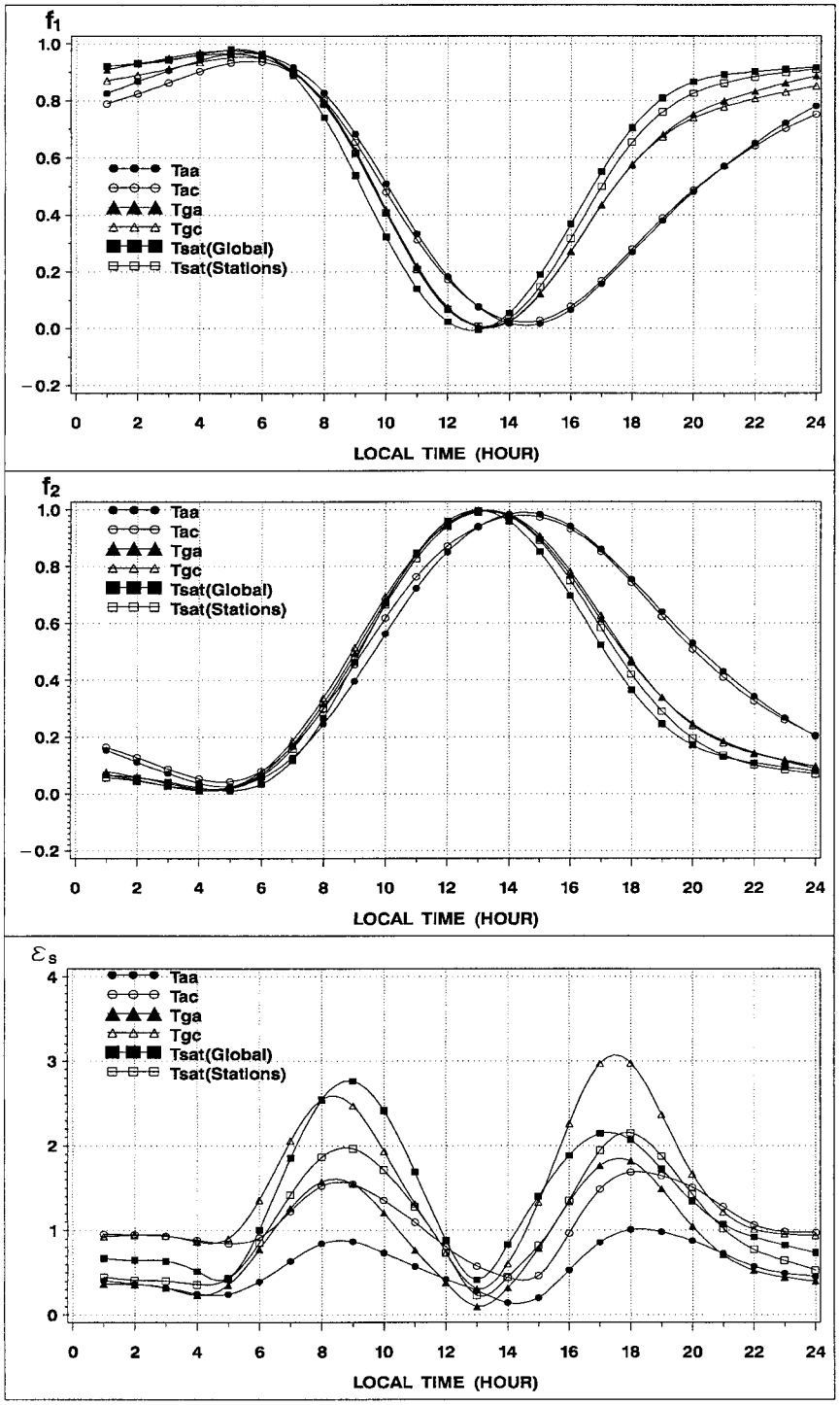

FIG. 5. The two basic functions $\mathbf{f}_{S,1}$ and $\mathbf{f}_{S,2}$ for approximating the MDCs using $T_{S,\max}$ and $T_{S,\min}$, and rmse of approximation.

analyzed using principal component analysis. Although the PCA statistics differ for different temperatures, for all of them the first two to three principal components—the loadings on the respective empirical orthogonal functions—accurately approximate the diurnal cycles.

The number of modes to keep in (1) depends upon the spatiotemporal resolution in data and accuracy requirements for a particular application. The rmse's are about 0.5° and 1°C for climatological (multiyear mean) air and ground/satellite temperatures, respectively, when

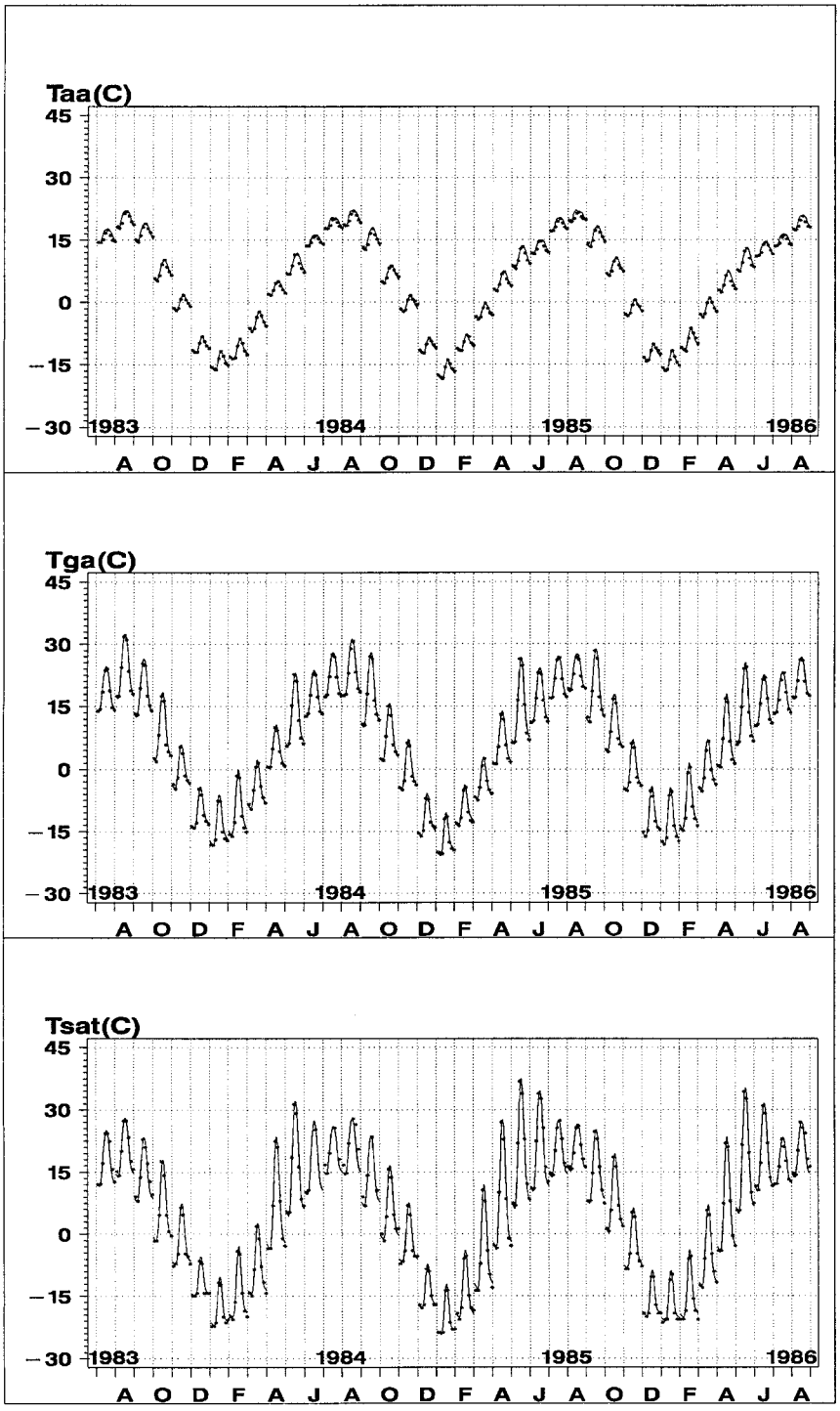

FIG. 6. Time series of T_{aa}, T_{ga}, and T_{sat} for July 1983–August 1986 at station Vladivostok: observations (dots) and their approximation using (4) through $T_{S,max}$ and $T_{S,min}$ (lines).

two modes are kept, and can be further lowered down by 20%–35% if the third mode is added. The two-mode approximation was shown to work satisfactorily when applied to data for individual years.

In any case, our study emphasizes the necessity of more than one variable to characterize the thermal regime at the surface. The practical implication is that mapping/monitoring of the surface temperature diurnal cycle should use two to three parameters. Each of them is presumably responsible for different geophysical pro-

cesses at the surface (for instance, the diurnal range is associated with soil moisture and vegetation state). Development of a comprehensive monitoring system and establishing respective relationships is the subject of future investigations. Instead of PCs, one can use a few temperatures, such as, for instance, daily maximum and minimum (or diurnal mean and range) in the two-mode approximation. In this case, station measurements of T_{min} and T_{max}, and data from NOAA afternoon satellites, collected once a day and once a night, can be potentially utilized to reconstruct the full MDC.

Out of four available station temperatures, the statistics of ground temperatures most closely resemble those of satellite data. Another important finding is that satellite retrievals, albeit made only under clear-sky conditions, do not reveal clear-sky bias at this level of analysis. Both results, not readily apparent intuitively, are obtained in this study purely empirically and therefore may be sample specific. The ranges and conditions of their validity need further verification. If confirmed, they provide an important guidance for validation of satellite retrievals using surface measurements and for the use of satellite data in conjunction with climate and weather prediction numerical models.

The algorithms for land surface temperature retrieval have undergone rapid development during recent years, suggesting that future satellite products will provide a better quality input to the analysis similar to the one carried out in this study. In the meantime, the present representation provides useful guidance for sampling and interpretation of future satellite products, and may be iteratively improved using new data. Thus, the ISCCP surface temperature data, which are by-products in the cloud project, remain a unique readily available source of temperature diurnal variability globally mapped in a convenient space–time grid. The monthly ISCCP land surface temperatures for individual years from 1983 to 1990 interpolated into 1-h local solar time grid by cubic splines are available in ASCII format through anonymous FTP (orbit-net.nesdis.noaa.gov. in directory /pub/crad/lst/ggutman/ts_land_isccp_interp).

Acknowledgments. We thank Drs. T. Karl and A. Basist, and two anonymous reviewers, for reviewing this paper and for valuable suggestions, particularly for pointing out the necessity of comparison with conventional temperature observations. We are greatly indebted to Dr. P. Groisman, NOAA/NCDC, for providing the fUSSR stations data and personal encouragement, and to Dr. J. Price for many useful recommendations and editorial improvements. Discussions with Dr. K. Vinnikov, University of Maryland, are appreciated. This work was done while A.I. was a University Corporation for Atmospheric Research visiting scientist at NOAA/NESDIS, on leave from the Marine Hydrophysics Institute, Sevastopol, Ukraine.

REFERENCES

Cairns, B., 1995: Diurnal variations of cloud from ISCCP data. *Atmos. Res.,* **37,** 133–146.

Cao, H. X., J. F. B. Mitchell, and J. R. Lavery, 1992: Simulate diurnal range and variability of surface temperature in a global climate model for present and doubled CO_2. *J. Climate,* **5,** 920–943.

Carlson, T. N., J. K. Dodd, C. G. Benjamin, and J. N. Cooper, 1981: Remote estimation of surface energy balance, moisture availability, and thermal inertia. *J. Appl. Meteor.,* **20,** 67–87.

Casellas, V., and J. Sobrino, 1989: Determination of frosts in orange groves from NOAA-9 AVHRR data. *Remote Sens. Environ.,* **29,** 135–146.

Cautenet, G., M. Legrand, and Y. Coulibaly, 1986: Computation of ground surface conduction heat flux by Fourier analysis of surface temperature. *J. Climate Appl. Meteor.,* **25,** 277–283.

Cracknell, A. P., and Y. Xue, 1996: Dynamic aspects of surface temperature from remotely-sensed data using advanced thermal inertia model. *Int. J. Remote Sens.,* **17,** 2517–2532.

Groisman, P., and P.-M. Zhai, 1995: Climate variability under clear skies: Applications for the cloud and snow cover feedback problems. *Proc. Sixth Int. Meeting on Statistical Climatology,* Galway, Ireland, All Ireland Committee on Statistics, 605–608.

——, E. Genichovich, R. Bradley, and B. Ilyin, 1997: Assessing surface–atmosphere interactions using former Soviet Union standard meteorological network data. Part II: Cloud and snow cover effects. *J. Climate,* **10,** 2184–2199.

Gutman, G., and A. Ignatov, 1995: Global land monitoring from AVHRR: Potential and limitations. *Int. J. Remote Sens.,* **16,** 2301–2309.

——, D. Tarpley, A. Ignatov, and S. Olson, 1995: The enhanced NOAA global land dataset from the Advanced Very High Resolution Radiometer. *Bull. Amer. Meteor. Soc.,* **76,** 1141–1156.

Idso, S. B., T. J. Schmugge, R. D. Jackson, and R. J. Reginato, 1975: The utility of surface temperature measurements for the remote sensing of surface soil water status. *J. Geophys. Res.,* **80,** 3044–3049.

ISCCP, 1992: *Monthly Cloud Products, July 1983–December 1990.* NASA, CD-ROM.

James, M. E., and S. N. V. Kalluri, 1994: The Pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring. *Int. J. Remote Sens.,* **15,** 3347–3363.

Karl, T. R., and Coauthors, 1993: Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.,* **74,** 1007–1023.

——, C. N. Williams Jr., P. J. Young, and W. M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States. *J. Climate Appl. Meteor.,* **25,** 145–160.

Kondragunta, C. R., and A. Gruber, 1994: Diurnal variation of the ISCCP cloudiness. *Geophys. Res. Lett.,* **21,** 2015–2018.

Lagouarde, J. P., and Y. Brunet, 1993: A simple model for estimating the daily upward longwave surface radiation fluxes from NOAA-AVHRR data. *Int. J. Remote Sens.,* **14,** 907–925.

Meinguet, J., 1979: Multivariate interpolation at arbitrary points made simple. *J. Appl. Math. Phys.,* **30,** 292–304.

Minnis, P., and E. F. Harrison, 1984: Diurnal variability of regional cloud and clear-sky radiative parameters derived from GOES data. Part I: Analysis method. *J. Climate Appl. Meteor.,* **23,** 993–1011.

Oke, T. R., 1987: *Boundary Layer Climates.* Cambridge University Press, 435 pp.

Preseindorfer, R. W., and C. Mobley, 1988: *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, 530 pp.

Price, J. C., 1977: Thermal mapping: A new view of the earth. *J. Geophys. Res.,* **82,** 2582–2590.

——, 1980: The potential for remotely sensed thermal infrared data to infer surface soil moisture and evaporation. *Water Resour. Res.,* **16,** 787–795.

Privette, J. L., C. Fowler, G. A. Wick, D. Baldwin, and W. J. Emery,

1995: Effects of orbital drift on Advanced Very High Resolution Radiometer products: Normalized difference vegetation index and sea surface temperature. *Remote Sens. Environ.,* **53,** 164–171.

Rossow, W., and R. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

——, and L. C. Garder, 1993: Validation of ISCCP cloud detections. *J. Climate,* **6,** 2341–2369.

Schaal, L. A., and R. F. Dale, 1977: Time of observation temperature bias and ''climatic change.'' *J. Appl. Meteor.,* **16,** 215–222.

Seguin, B., D. Courault, and M. Guerif, 1994: Surface temperature and evapotranspiration: Application of local scale methods to regional scales using satellite data. *Remote Sens. Environ.,* **49,** 287–295.

Tarpley, D., 1994: Monthly evapotranspiration from satellite and conventional meteorological observations. *J. Climate,* **7,** 704–713.

Wetzel, P. J., D. Atlas, and R. H. Woodward, 1984: Determining soil moisture from geosynchronous satellite infrared data: A feasibility study. *J. Climate Appl. Meteor.,* **23,** 375–391.